Universiteit Utrecht

NCRM National Centre for Research Methods

Investigation of Nonresponse Bias and Representativeness in the First Cross-National Probability Based Online Panel (CRONOS)

Olga Maslovskaya (University of Southampton)
Peter Lugtig (Utrecht University)
Gabriele Durrant (University of Southampton)

17 July 2019, ESRA Conference 2019, Zagreb

E·S·R·C ECONOMIC & SOCIAL RESEARCH COUNCIL

UNIVERSITY OF Southampton

MANCHESTER 1824 The University of Manchester

THE UNIVERSITY OF EDINBURGH

Funding

ESRC Secondary Data Analysis Initiative Project "Understanding survey response behaviour in a digital age: Mixed-device online surveys and mobile device use"

Motivation

- Introduction of online probability based panels such as NatCen panel and CRONOS panel
- Approach to invite sample members: On the back of a probability-based face-to-face survey but potential problems with attrition and resulting representativeness
- Not much is known about representativeness in CRONOS yet

Main Aim

- To investigate representativeness across waves and between countries in CRONOS online cross-national probability-based panel

CRONOS (Cross-National Probability Based Online Panel)

- Collected on the back of the ESS Round 8

Countries:

  - Estonia (EE)
  - Great Britain (GB)
  - Slovenia (SI)

- After completing the ESS face-to-face interview, those 18+ were invited to participate in a 10-minute welcome survey and six 20-minute online surveys over a time period of 12 months
- Data collection took place between December 2016 and February 2018

Country contexts

- Internet penetration - 2018
  - UK – 89.8%
  - Estonia – 80.0%
  - Slovenia – 70.7%

Stages of data collection in CRONOS (1)

1	Gross sample - ESS	
2	Responded to ESS R8	
3	Those who were expected to participate	No 15-17 year olds No NI in GB
4	Initial panel invitation	No clear why there are differences between 3 and 4
5	Recruitment interview	Responded to initial panel invitation: said yes or unsure to participation
6	Wave 0	Not all panellists were invited to Wave 0 as it was launched in December 2016 when ESS fieldwork was still in progress
7	Wave 1	
8	Wave 2	
9	Wave 3	
10	Wave 4	
11	Wave 5	
12	Wave 6	

Two separate parts of the process

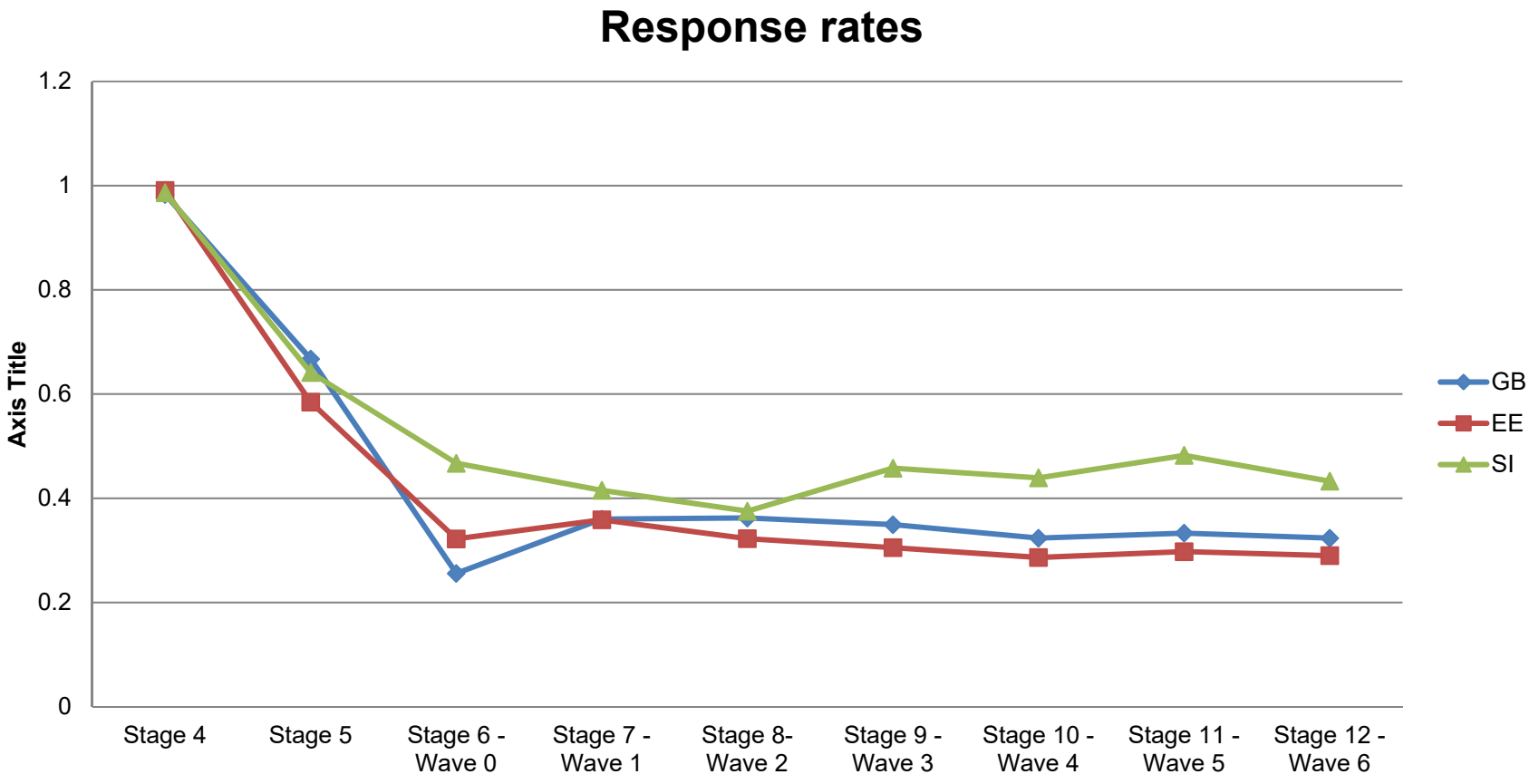
1. Stages 1 and 2 – not presented here
2. Stages 3-12 (some groups completely removed from the original sample (15-17 year olds and NI in GB) – stage 3 is used as base for fitting response models and for calculations of representativeness indicators

NCRM National Centre for Research Methods

Stages of data collection in CRONOS (2)

		GB	Estonia	Slovenia
1	Gross sample	4,447	2,901	2,278
2	Responded to ESS R8	1,885	2,018	1,303
3	Those who were expected to participate (no 15-17 + no NI in GB)	1,786	1,962	1,238
4	Panel invitation	1,756	1,944	1,235
5	Recruitment interview: responded to initial panel invitation	1,192	1,147	803
6	Wave 0	457	632	585
7	Wave 1	643	704	520
8	Wave 2	647	633	470
9	Wave 3	624	599	574
10	Wave 4	578	562	550
11	Wave 5	595	584	604
12	Wave 6	578	569	542

Response rates



Methods

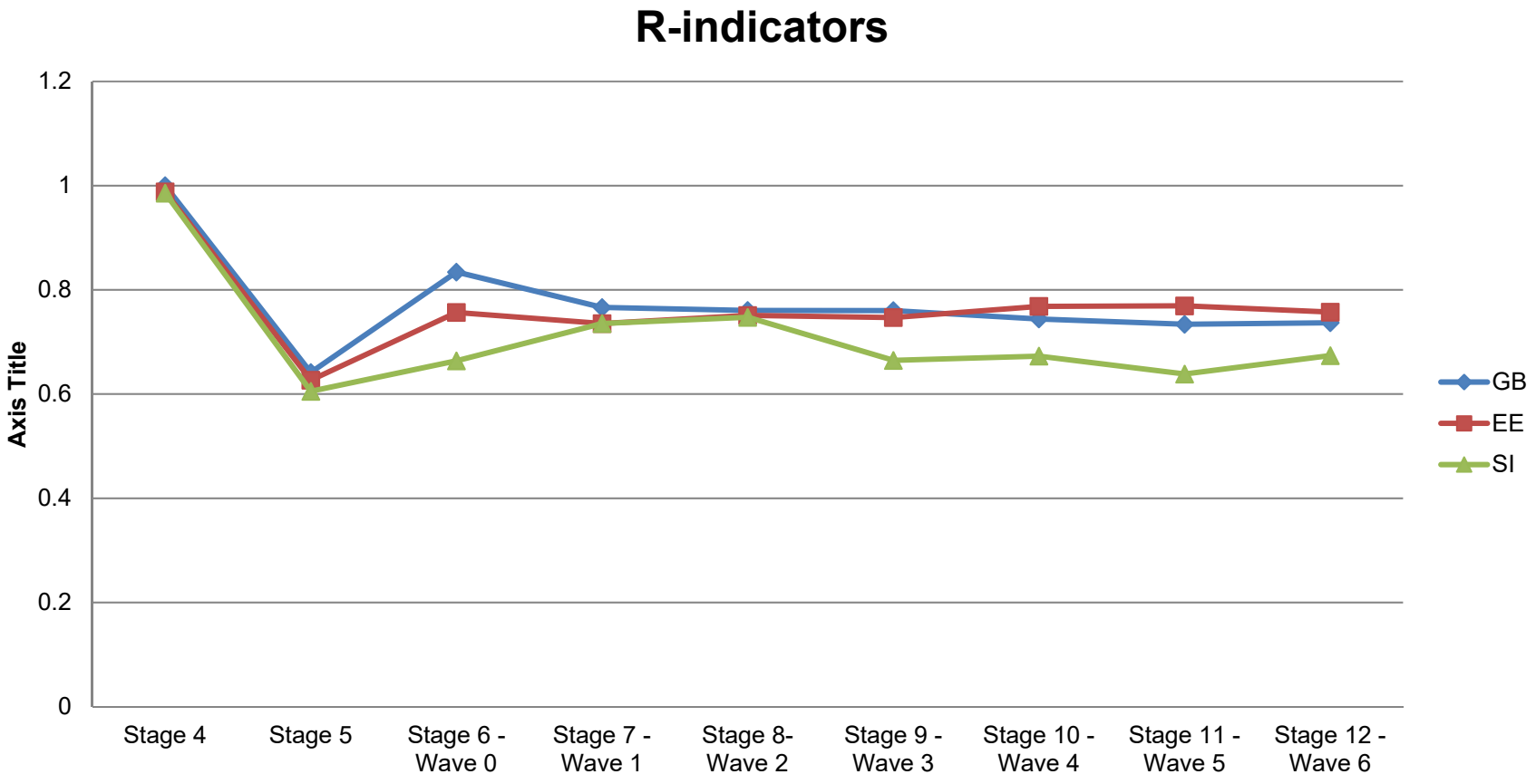
- **R-indicators** – representativeness indicator (Schouten et al. 2009; Schouten et al. 2011; de Heij et al. 2015) – designed to measure the degree to which the respondents to a survey resemble the complete sample
- Large indicators imply representativeness (0:1)
- R-indicators do not provide means to identify subgroups for targeting and prioritising

- **Partial R-indicators** – designed to evaluate the contribution of a single specified auxiliary variable or category within a variable to a lack of representative response

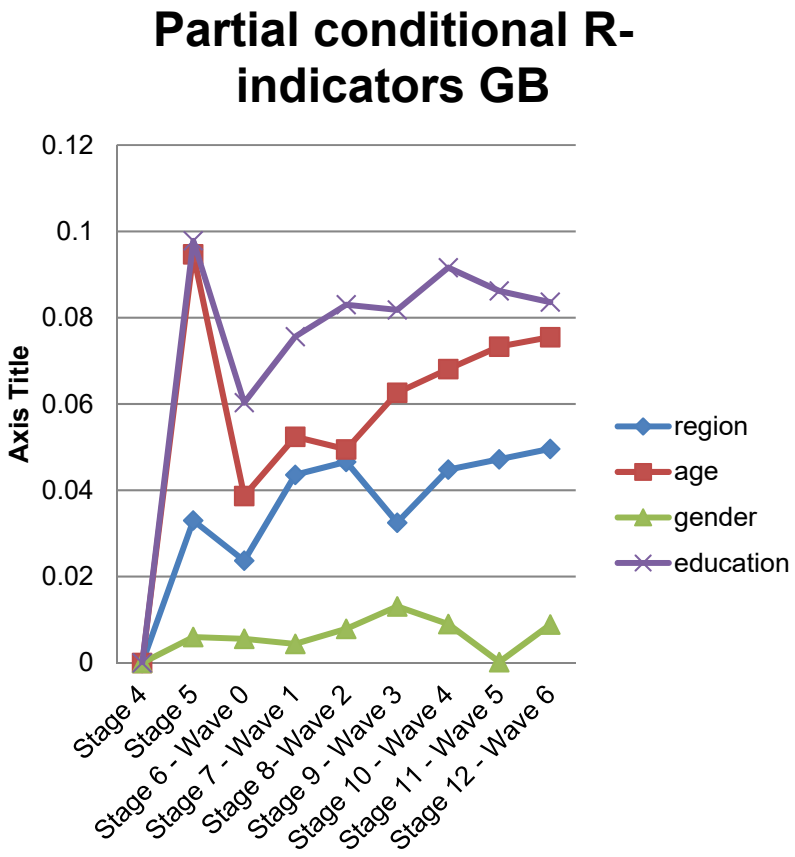
- **Conditional partial R-indicators on the variable level** – the larger the indicator, the higher the role of this variable in reducing the lack of representativity as it has the higher impact on bias
- **Conditional partial R-indicators within categories** – the larger the indicator, the larger the impact of that category on nonresponse bias

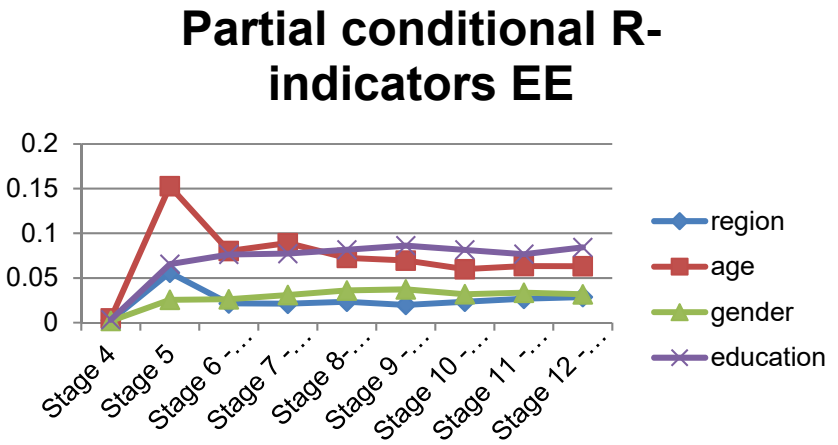
- Standard **response propensity models** (logistic regressions) with control variables: age, gender, education, region

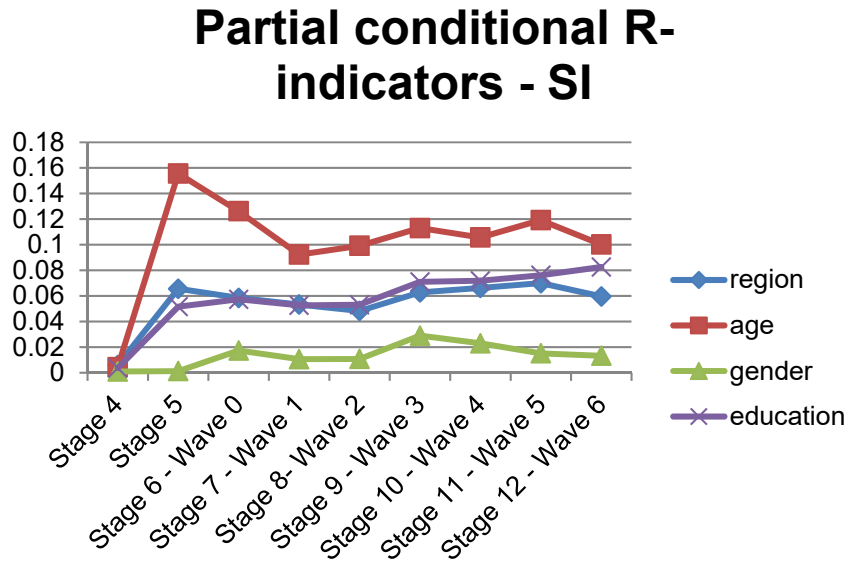
Results (1): R-indicators



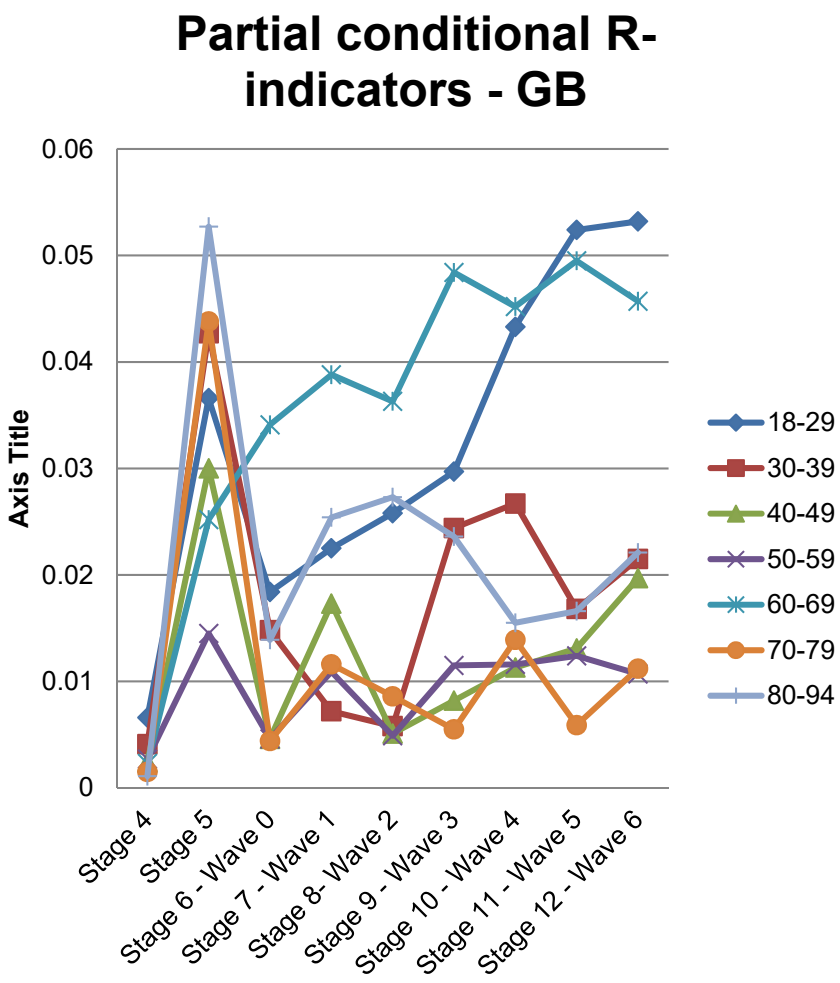
Results (2): Partial conditional R-indicators

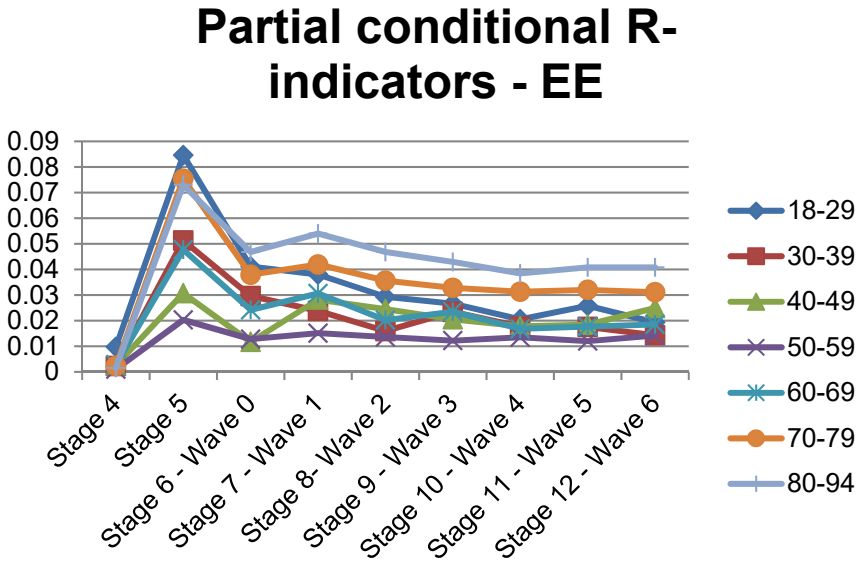


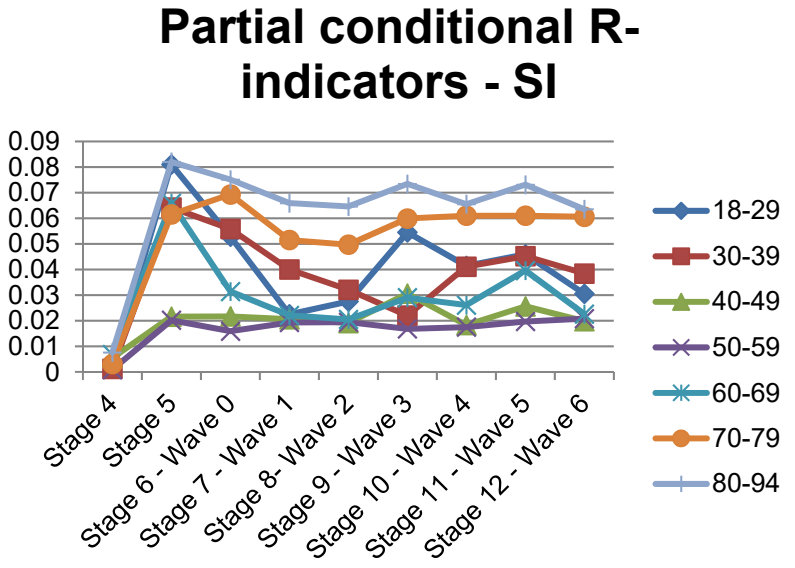




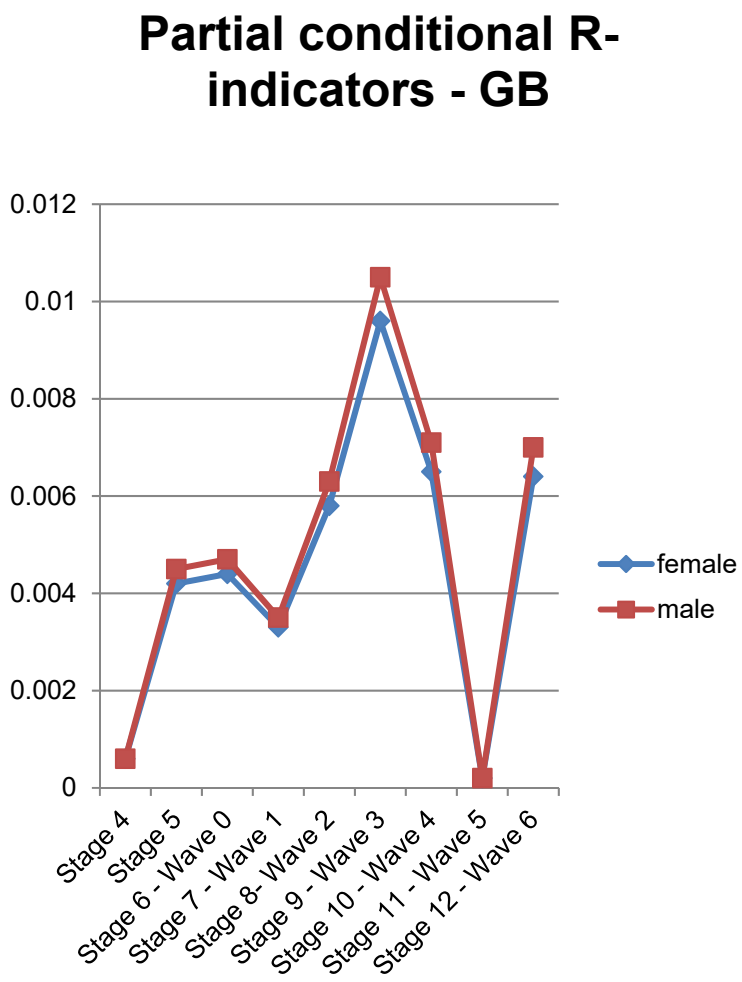
Results (3): Age

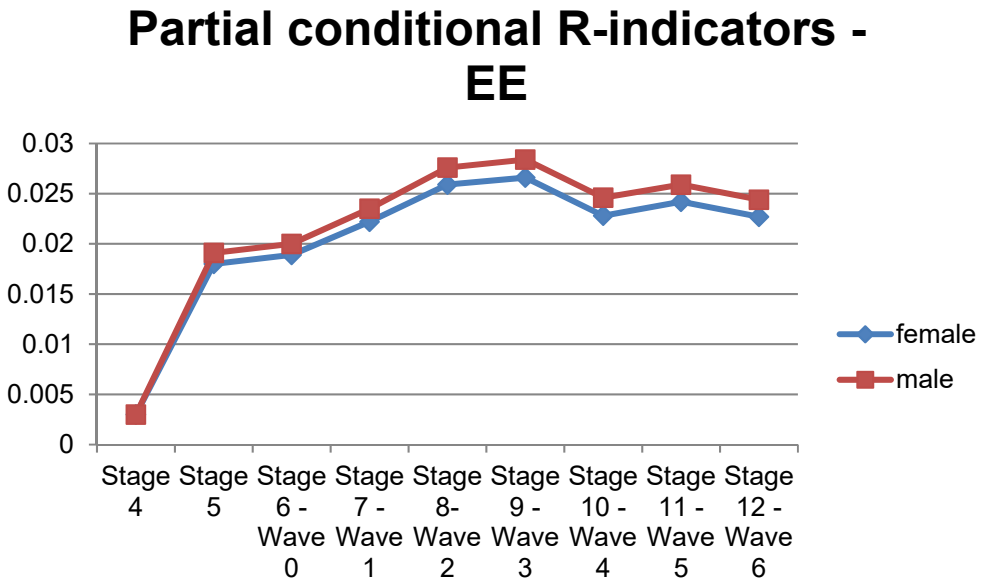


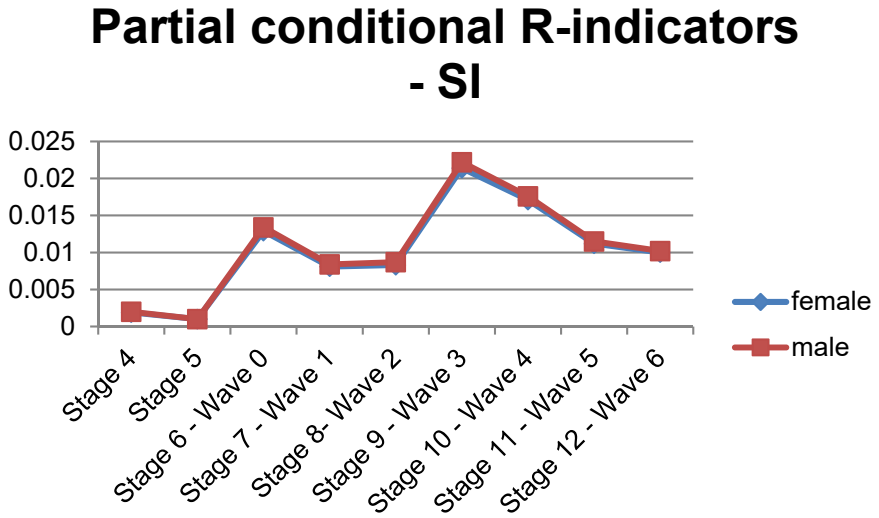




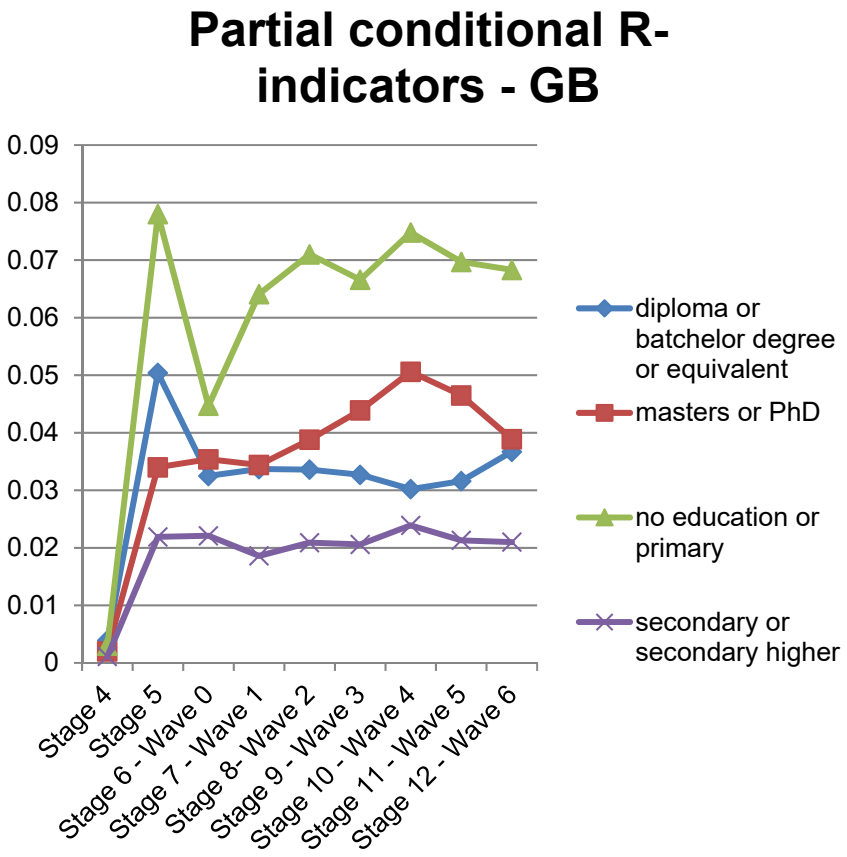
Results (4): Gender

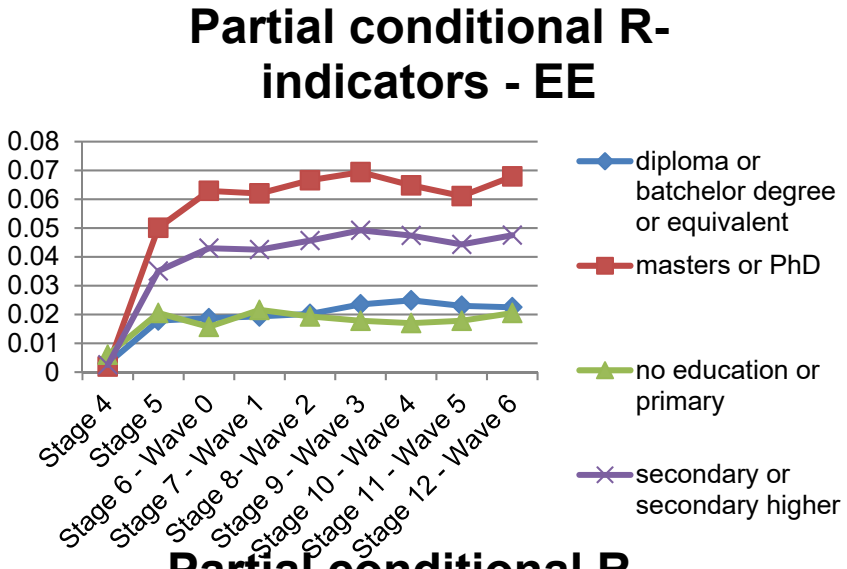


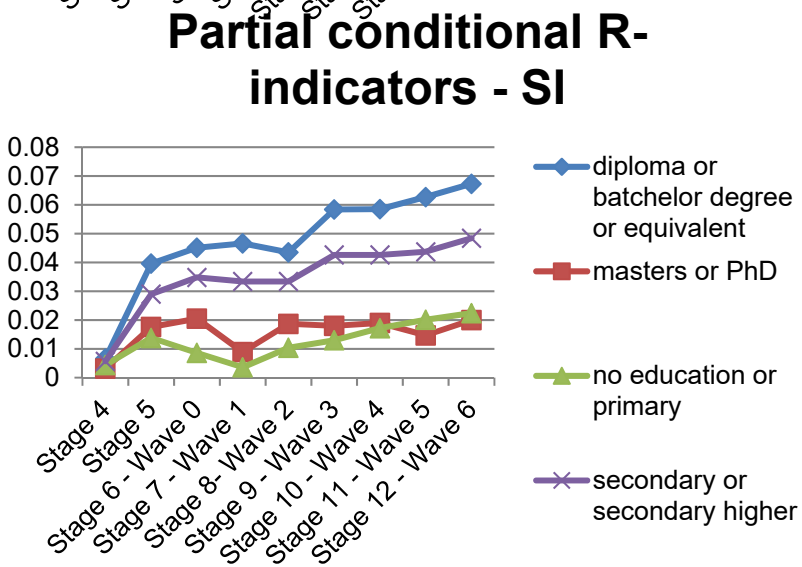




Results (5): Education







Limitation

- R-indicators can overestimate representativeness in specific contexts (low response rates and low propensity variation). Coefficients of variation (CVs) were also used but not presented here.

- Absence of paradata for device used by respondents to complete survey for Wave 1 and Wave 2
- Would be very important to obtain these paradata for CRONOS 2 as would be very important to investigate differences by device

Universiteit Utrecht

NCRM National Centre for Research Methods

Conclusions

- Important to look at R-indicators to have more complete picture of data quality
- Differences in repersentativeness across countries are relatively small (response rates are different across countries). Also differences are small across time within countries
- Partial R-indicators show country-specific contexts and are very useful for targeting and prioritising of specific subgroups
- Results for the UK are consistent with results obtained from other social surveys in the UK, including Understanding Society

References

de Heij, V., Schouten, B., Shlomo, N. (2015). *RISQ manual 2.1: Tools in SAS and R for the computation of R-indicators, partial R-indicators and partial coefficients of variation*. Resresentativity Indicators for Survey Quality (RISQ) project. http://hummedia.manchester.ac.uk/institutes/cmist/risq/RISQ-manual-v21.pdf

Moore, J. C., Durrant, G. B., & Smith, P. W. F. (2016). Data set representativeness during data collection in three UK social surveys: Generalizability and the effects of auxiliary covariate choice. *Journal of the Royal Statistical Society: Series A (Statistics in Society) 181(1): 229-248*.

Moore, J., Durrant, G., & Smith, P. W. F. (2017). Correlates of record linkage and estimating risks of non-linkage biases in business datasets. *Journal of the Royal Statistical Society: Series A (Statistics in Society) 181 (4): 1211-1230*.

Schouten, B., Cobben, F., Bethlehem, J. (2009). Indicators for the representativeness of survey response. *Survey Methodology* 35: 101-113.

Schouten, B., Shlomo, N, Skinner, C. (2011). Indicators for Monitoring and Improving Representativeness of Response. *Journal of Official Statistics* 27 (2): 231-253.

Universiteit Utrecht

NCRM National Centre for Research Methods

Thank you very much!

om206@soton.ac.uk